社会奖惩对欺骗行为的影响*

袁　博　赵靖实　漆　丹　赵　彤　胡佳琪

(宁波大学心理学系暨研究所, 浙江 宁波 315211)

摘 要　欺骗是指通过言语或非言语方式向他人提供错误信息或隐瞒相关信息的行为，奖励和惩罚是影响欺骗行为的关键因素。然而，以往研究主要关注金钱奖惩，对于社会奖惩如何影响欺骗行为目前尚不清楚。本研究通过 3 项实验，探讨了社会奖惩对欺骗行为的影响及其中介与调节机制。实验 1 采用信号博弈任务考察社会奖惩能否对欺骗行为产生影响。结果发现，与金钱奖惩类似，社会奖惩能够减少欺骗行为，且社会惩罚比社会奖励的效果更好。采用漂移扩散模型分析发现，社会奖惩与金钱奖惩条件下漂移率v显著小于无奖惩条件，表明社会奖惩与金钱奖惩促使个体更倾向于累积选择不欺骗的证据。实验 2 和 3 采用相同的实验任务，进一步考察了声誉关注在社会奖惩影响欺骗行为中的中介作用，以及社会价值取向的调节作用。结果发现，声誉关注在社会奖惩对欺骗行为的影响中起中介作用，社会价值取向调节社会奖惩通过声誉关注影响欺骗行为的中介过程。上述结果表明，社会奖惩能够作为抑制欺骗行为的有效手段。

关键词　社会奖惩, 欺骗, 声誉关注, 社会价值取向, 漂移扩散模型
分类号　B849: C91

1 引言

欺骗(deception)是通过言语或非言语的方式向他人提供错误信息或隐瞒相关信息以达到某种目的的行为(Depaulo et al., 2003)，它涉及一系列认知和行为过程，包括信息的选择、加工和传递等(Carr et al., 2019)。欺骗在社会生活中普遍存在，对组织和社会都造成了巨大的代价(Mazar & Ariely, 2006)。为了更好地制定减少欺骗行为的政策和制度，我们需要了解哪些因素影响人们的欺骗行为。人们在什么情况下会选择欺骗？根据标准经济模型中的理性人假设：个体在进行欺骗行为时，会权衡欺骗行为所带来的预期外部收益和成本(Allingham & Sandmo, 1972; Becker, 1968)。因此，奖励与惩罚(reward and punishment)作为增加或减少某

收稿日期: 2023-11-22

* 全国教育科学规划一般项目“青少年道德决策中的同伴影响及其认知情感机制研究” (BBA210033)。

通讯作者：袁博, E-mail: yuanbopsy@gmail.com

种动机或行为的方式，也是干预欺骗行为的重要手段。

已有大量研究探讨了金钱等物质形式的奖惩对欺骗行为的影响(Gneezy, 2005; Kaushik et al, 2022)。如，Mazar 等人(2008)发现，当欺骗可以获得更多的利益时，个体会表现出更多的欺骗行为。而当诚实行为受到金钱奖励时，个体的欺骗行为会减少(Rosenbaum et al., 2014)。同样，金钱惩罚通过增加欺骗成本也抑制了欺骗行为。研究发现，增加对欺骗行为金钱惩罚的可能性与力度均能够减少个体的欺骗行为(Behnk et al., 2018; Nagin & Pogarsky, 2003)。然而，金钱奖惩是一把双刃剑，在抑制欺骗行为的同时，也带来了高昂的成本。此外，个体一旦适应了外部的金钱强化，减少或者结束这种强化可能导致期望的行为退回到基线甚至更低水平(Mulder et al., 2006)。比如，金钱惩罚会将合作、信任等亲社会行为的内在动机转变为外在动机，在惩罚撤除后反而会降低亲社会行为(Gächter & Herrmann, 2009)。金钱奖惩还可能使个体将欺骗行为视为一种可接受的风险。当欺骗带来的收益超过金钱惩罚时，个体可能会认为欺骗是合理的，从而在某种程度上鼓励欺骗行为。是否存在一种更具潜力的奖惩方式抑制欺骗行为？

1.1 社会奖惩的概念及其操纵形式

人类具有社会属性，从简单的言语交流到个体间的利益交换，都需要与他人进行互动，欺骗行为通常也发生在与他人的互动中(Kohls et al., 2013)。社会奖励(social reward)是在没有物质回报的情况下，个体所期望获得的具有社会性质的结果，包括尊重、礼貌、接纳、称赞或认可等行为和言语表达(Ramirez-Marin & Shafa, 2018)。社会奖励意味着个体被他人或群体的接受和喜爱，能够带来满意、快乐、自豪等积极情绪(De Cremer & Tyler, 2005)。社会惩罚(social punishment)是指通过批评、排斥、谩骂、流言蜚语等形式进行的惩罚(Kim & Jeong, 2020)。社会惩罚会引起个体的负面情绪，如失落、沮丧、焦虑、嫉妒、抑郁等，还可能会激活类似于躯体伤害的疼痛反应的神经系统反应(Beston, 2019)。

以往研究采用不同的方式操纵社会奖惩。例如，有研究采用笑脸或悲伤脸代表社会奖惩(Wang et al., 2017)，也有研究使用了向上和向下的箭头代表社会奖惩(Wang et al., 2020)，还有研究使用不同程度的快乐面部表情代表社会奖励(Spreckelmeyer et al., 2009)。表情符号(Emoji)以一种独立的表达方式来传达意义(Kaye et al., 2017)，有效地发挥了类似传统面对面互动中非语言线索的功能，能够传达丰富的社会与情感信息(Boutet et al., 2021; Cherbonnier & Michinov, 2021; Fischer & Herbert, 2021; Hand et al., 2023)。根据社会信息模型(Walther, 1992)，从表情符号中识别的情绪可以引导个体随后的社会行为(Van Kleef, 2009)。已有研究表明，表情符号提供的线索有助于个体表达特定的情感状态，并强化沟通信息中所蕴含的情

感内容(Kaye et al., 2016)。表情符号可以模拟面对面交流中的表情反馈，提供积极或消极的评价信息，从而影响信息接收者的情绪反应和对信息的理解(靳宇倡 等, 2022; Kaye et al., 2016)。那么，基于表情符号传达的社会奖惩能否影响欺骗行为？这种社会奖惩影响欺骗行为可能的心理机制是什么？

1.2 社会奖惩的对欺骗行为的影响及其中介机制

社会奖惩是人类互动行为中的一个关键激励因素，为了获得社会奖励或者避免社会惩罚，人们可能会放弃或牺牲自己的金钱利益，以争取融入某个社会群体或获得与他人互动的机会(Tamir & Mitchell, 2012; Tamir et al., 2015)。保持良好的社会形象符合个人的长期利益(Gintis, 2000)，那些成功展现慷慨行为的人在声誉和地位上通常会获得积极回报(Flynn et al., 2006; Hardy & Van Vugt, 2006)。此外，人们还会运用多种人际沟通策略来管理自身形象，从而获得他人的接纳、认可和好感(Jones & Pittman, 1982; Berman et al., 2015)。因此，满足社交、尊重等社会性需要与满足物质需求一样，对于人类生存至关重要(Romano et al., 2017)。同时，为了避免社会排斥或声誉损害等形式的社会惩罚，个体往往愿意在公共物品博弈(public good game, PGG)中付出更多金钱(Feinberg et al., 2014; Guala, 2012)。社会排斥的威胁是促进合作的一种重要方式。研究发现，在公共物品博弈中，当低贡献者在被排斥前收到警告时，大多数低贡献者会通过更高的贡献回应警告，以避免被排斥(Cinyabuguma et al., 2005)。上述研究结果表明，个体具有强烈的追求社会奖励和避免社会惩罚的动机。因此，**我们推测社会奖惩也能作为抑制欺骗行为的有效手段，减少个体在人际互动中的欺骗行为**。

如果社会奖惩能够减少欺骗行为，那么其中可能的心理机制是什么？Mazar 等人(2008)提出的自我概念维持理论(theory of self-concept maintenance)认为，个体在进行欺骗行为时会面临心理冲突(欺骗获利 vs.维持积极自我概念)。为了解决这一冲突，个体会通过调整欺骗行为的程度，尽可能避免对自我概念(如，声誉)造成过度威胁。声誉是一种通过观察个体历史行为而形成的总体印象，是需要持续投资和维护的无形资产(Pfeiffer et al., 2012)。声誉关注(reputation concern)，即个体对自身声誉的关注程度，是维持积极自我概念的重要方式。在涉及声誉的情境下，个体通常会调整行为以符合社会规范，从而避免负面评价并维持良好的自我概念(Leary & Kowalski, 1990)。

已有研究表明，声誉关注是影响个体是否选择欺骗的重要心理因素(Russell et al., 2008)。声誉关注的前提是社会评价的可能性，相较于金钱奖惩所强调的经济权衡，社会奖惩涉及更多的社会评价的过程(Deci, 1971)。来自他人的社会奖励可能提高个体的声誉，而来自他人的社会惩罚则会损害个体的声誉。例如，研究发现，公开捐献者的名字以提升其声誉关注，能

够显著增加慈善捐款的数额(Karlan & Mcconnell, 2014)。Vabba 等人(2022)发现，被试通常会根据声誉受损的风险程度调整自己的行为，当他们意识到自己的行为可能被观察到时，欺骗行为明显减少。类似的，Ariely 和 Gino(2012)在研究群体环境中的自我信号与声誉关注对不诚实行为的影响时发现，当个体的道德标准被提醒或处于被他人观察和评价的情境中时，他们的欺骗行为显著减少。综上所述，社会奖惩可能通过增强个体的声誉关注减少欺骗行为。具体而言，社会奖励可能增强个体的积极声誉，从而减弱其欺骗动机；而社会惩罚则通过威胁个体声誉促使他们回归规范行为。因此，**我们推测声誉关注在社会奖惩对欺骗行为的影响中起到中介作用**。

1.3 社会价值取向在社会奖惩影响欺骗行为中的调节作用

社会奖惩影响欺骗行为可能受到人格因素的调节，社会价值取向(social value orientation, SVO)作为一种稳定的社会偏好，反映了个体在自身和他人之间分配利益的倾向(Balliet et al., 2009; Grosch & Rau, 2017; Steinel, 2015)。社会价值取向可以简化为两类：(1)亲自我取向(proself)：包括个人倾向和竞争倾向，强调自身的利益高于他人利益；(2)亲社会取向(prosocial)：包括合作和平等取向，关注自己与他人的共同利益，甚至更关注他人的利益(张振 等, 2015)。已有研究表明，社会价值取向对奖惩反应存在显著差异。例如，Balliet 和 Van Lange(2013)的研究发现，相较于亲社会取向者，金钱奖励对亲自我取向者的合作行为具有更强的激励效果。而亲社会取向个体则对与社会接纳和社会拒绝相关的社会奖惩表现出更高的敏感性(李珺 等, 2020)。这种对社会奖惩的敏感性，可能使亲社会取向的个体面对社会奖惩时更关注自身声誉，从而影响他们的欺骗行为。声誉作为一种“社会货币”，对重视社会关系的亲社会取向者尤为重要(Milinski et al., 2002)。亲社会取向个体通常更重视与他人的关系，倾向于通过维护积极声誉促进信任与合作，他们在互动中更关注自己的声誉，以确保他人对自己的积极评价(Simpson & Willer, 2008)。相比之下，亲自我取向个体倾向于最大化自身利益，而非建立长期关系。在短期获益的情境下，亲自我取向者更容易忽略声誉受损的风险，对声誉的关注相对较低(Van Lange, 1999)。基于此，我们推测，**社会价值取向在社会奖惩通过声誉关注影响欺骗行为的过程中起到调节作用**。具体而言，与亲自我取向个体相比，社会奖惩更容易引发亲社会取向个体对声誉的关注，从而减少他们的欺骗行为。

综上，本研究拟通过 3 项实验探讨社会奖惩对欺骗行为的影响及其心理机制。实验 1 考察社会奖惩能否对欺骗行为产生影响，实验 2 进一步考察声誉关注在社会奖惩对欺骗行为的影响中的中介作用，实验 3 考察社会价值取向在上述过程中的调节作用。

2 实验 1 社会奖惩对欺骗行为的影响

2.1 被试

基于本研究的实验设计(单因素被试内)，设定显著性水平 $\alpha = 0.05$，统计检验力 $1\text{-}\beta = 0.8$，达到中等效应量 $\eta_p^2 = 0.13$。根据 Gpower 3.1 软件计算，样本量至少需要 28 名被试。因此，招募了 30 名在校大学生，平均年龄为 20.43 ($SD = 1.63$)，其中男生 10 人，女生 20 人。被试实验前签署知情同意书，实验后给予一定的实验报酬。

2.2 实验设计

采用单因素被试内实验设计，自变量为奖惩类型，分为社会奖惩、金钱奖惩、无奖惩三个水平；因变量为欺骗行为，即被试在信号博弈任务中发送虚假信息的比例。

2.3 实验任务

采用信号博弈任务(the signaling game)测量被试欺骗行为。该任务包括信息发送者与信息接收者两个角色；其中信息发送者拥有信息，而信息接收者只能听取发送者的信息。该任务创建了一个信息不对称情境，体现了欺骗行为中的核心困境，即信息发送者与信息接收者的利益冲突。发送者若选择如实发送信息，则会减少自己的收益而给接收者带来更大收益；若选择发送虚假信息可能会增加自己的收益，而减少接收者的收益(Zhu et al., 2014)。

被试始终充当游戏中的信息发送者，任务开始后向被试展示四个选项，每个选项对应不同的金钱收益(例如，选项 1 对应信息发送者获得 10 元，信息接收者获得 7 元，而选项 3 对应发送者 7 元，接收者 10 元，选项 2、4 为零收益选项，见图 1)。设置两个零收益选项是为了让被试相信就算自己选择欺骗对方，对方出于对自身收益的考虑，也会根据被试发送的信息进行选择。否则就有 50%的可能性选中零收益选项，导致双方收益为 0，造成更大的损失，从而让被试相信对方会始终根据自己发送的信息进行选择。被试需向接收者推荐其中一个选项，并通过发送信息告诉对方这个选项会使其收益更高。被试可以选择发送真实信息，以帮助接收者获取更高的收益；也选择欺骗对方，使自己获得更高的收益。告知被试系统有一定概率会向对方显示其所发送信息的真实性，然后对方会根据信息的真实性对被试进行奖励或者惩罚。若系统未向对方显示信息的真实性，对方将无法得知被试所发送信息是否真实，即无奖惩反馈。

参照 Gneezy(2005)的实验设计，被试通过欺骗能够为自身带来的利益分为高获利、中获利、低获利三个水平，给信息接收者造成的利益损失也分为高损失、中损失、低损失三个水平，其中获利或损失金额低于 5 元为低水平，5~10 元之间为中水平，高于 10 元为高水平。

关于社会奖惩的操控，以往的研究中通常使用认可或者不认可的评语，大拇指向上或向下的手势，开心或者沮丧的人脸或者卡通图片等方式(Matyjek et al., 2020)。本研究参照了 Wang 等人(2017)的方法，采用开心或者沮丧的表情符号作为社会奖惩。金钱奖惩则采用了“+”或“–”硬币图片的方式操控，且金钱奖惩的金额设置为一元，其金额低于欺骗所获得的利益，以防止被试选择诚实是为了更大的获利。

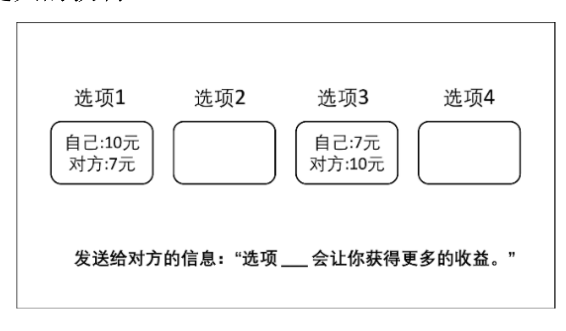


图 1 信号博弈任务示意图

2.4 实验流程

实验任务分为社会奖惩、金钱奖惩与无奖惩三个区块(block)，每个区块 40 个试次(trial)，三个区块的顺序在被试间随机以平衡顺序效应。被试将始终充当游戏中的信息发送者，并被告知信息接收者为随机匹配的另一名被试，在隔壁实验室房间与之进行互动，实则为计算机设定的程序。在每个阶段都会告知被试系统会有一定的概率向对方显示其所发送信息的真实性，实际设置向对方显示信息真实性的概率为 80%，但具体概率被试并不知道。若该试次系统向对方显示了信息的真实性，对方则会根据信息的真实性给予被试奖励或惩罚，若系统未向对方显示信息是否真实，则无奖惩直接进入下一个试次。在不同区块中被试所收到的奖惩反馈不同，在社会奖惩阶段，被试会收到开心或者沮丧的脸；在金钱奖惩阶段，被试会收到增加(“+”)或者扣除(“–”)一元硬币；无奖惩阶段则无论系统是否向对方显示信息真实性，对方都不会给被试奖惩。此外，在每一个区块中，有 20%的试次中还对被试的情绪进行了测量(对方的回应让我感到高兴的程度，请按键盘上对应的数字进行选择：*1 非常不高兴~9 非常高兴*)[1]。情绪测量试次采用伪随机的方式，在每个区块(40 个试次)中选择了固定的 8 个信息公开的试次(20%)进行情绪测量。

每个阶段的具体任务流程如下：首先，屏幕上会呈现一个“+”，呈现时间为 800ms，提示本轮实验任务开始。600ms 的空白屏后会呈现信息发送的四个选项，并提示被试选择发送

[1] 为了确保奖惩引发的情绪不会影响我们的研究结论，我们在实验设计中采取了伪随机的方法对情绪进行测量，这一处理方法能够有效排除情绪对研究结果的潜在干扰。

信息。被试做出选择后，屏幕上将会将被试所选的信息加亮以确认被试所做出的选择(1200ms)。之后，会提示被试系统是否向对方显示了信息的真实性，持续时间在 1000ms ~ 2500ms 之间随机，以模拟对方做出奖励或惩罚的选择。若该试次系统向对方显示了信息的真实性，且被试所发送信息为真，则给予奖励，反之则给予惩罚，奖惩反馈呈现时间为 1500ms。实验结束后，计算机将在每个阶段中随机选择一个试次，将根据被试在该试次的选择及该试次中的奖励或惩罚值获得报酬。图 2 以社会奖惩区块为例显示了实验流程，实验任务采用 PsychoPy 软件(Peirce, 2009)在计算机上呈现。

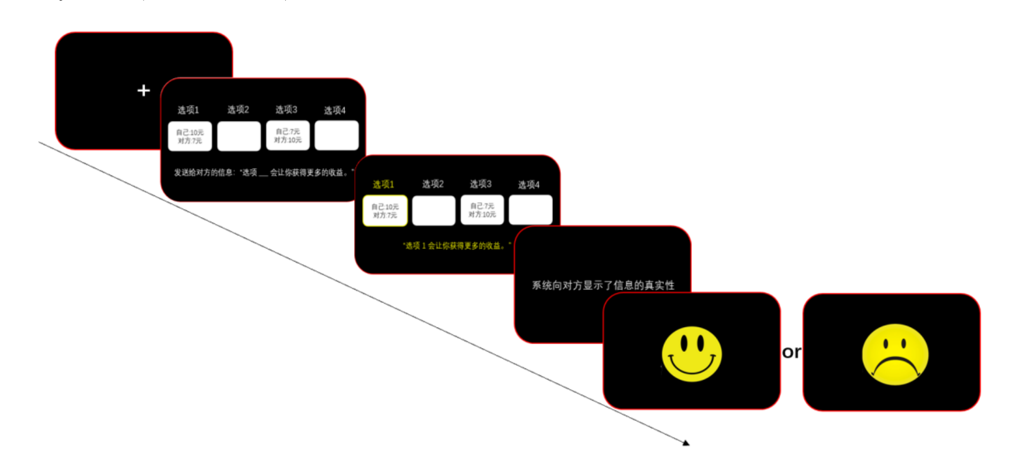


图 2 实验 1 任务流程图

2.5 数据分析处理

对于反应时数据，我们首先进行了传统的数据分析。然后，采用漂移扩散模型(drift-diffusion modeling, DDM)对反应时数据进行分析。DDM 把决策描述为一个连续的抽样过程，即带有噪声的信息从起点累积到对应于某一选项的边界或阈值，该选项被选中(Ratcliff & McKoon, 2008)。DDM 模型参数包括漂移率(drift rate, ν)、边界高度(boundary, α)、起始点偏差(bias, z)和非决策时间(non-decision-time, τ)。其中，漂移率ν代表累积某一选择证据的速率，即个体倾向于某一选项的偏好越强烈，信息向该选项积累的速度就越快；边界高度α表示在做出反应之前必须积累的信息量；起始点偏差z表示决策之前的先验偏向；非决策时间τ反映了影响决策反应时中的其他因素，包括信息编码与按键反应的时间(袁博 等, 2023; 张银花 等, 2020)。在本研究中，DDM 使用选择和反应时分布来描述被试如何累积证据做出欺骗或不欺骗选择。漂移率ν量化了被试通过加工选项收益信息获得的有利于选择欺骗或不欺骗证据的强度，即对选择欺骗或不欺骗的价值权衡程度。在本研究中，我们将选择欺骗编码为 1，选择不欺骗编码为 0。因此，漂移率ν越正，越偏向于选择欺骗；漂移率ν越负，越偏向于选择不欺骗。起始点偏差z量化了被试在获取任何证据之前选择欺骗/不欺骗的倾向。

边界高度α量化了被试在做出选择时需要的证据量，进而反映了不同条件下个体做出选择的谨慎程度。

采用基于 Python 的软件包 HSSM (Hierarchical Sequential Sampling Modeling)对基于反应时的 DDM 模型进行贝叶斯参数估计。贝叶斯参数估计可以直接对估计出的后验(posterior)参数进行差异性检验，而不用依赖于传统的频率统计分析。选取两个感兴趣条件下参数的后验分布(如社会奖惩和无奖惩条件下的漂移率)，如果这两个条件下的后验分布的 95%的最高密度区间(highest density interval, HDI)没有相互重叠，那么两个条件之间后验分布的差异就是可信的(袁博 等, 2023)。

2.6 结果

首先，对社会奖惩、金钱奖惩、无奖惩三种条件下被试的情绪水平进行重复测量方差分析。结果发现，三种条件下的被试的情绪无显著差异，$F(2, 58) = 2.99, p = 0.058, \eta_p^2 = 0.09$。其次，对社会奖惩、金钱奖惩、无奖惩三种条件下被试的欺骗比例进行重复测量方差分析。结果发现，被试在三种条件下的欺骗比例存在显著差异，$F(2, 58) = 12.66, p < 0.001, \eta_p^2 = 0.30$。其中社会奖惩($M = 0.34, SD = 0.27$)和金钱奖惩($M = 0.34, SD = 0.25$)条件下的欺骗率显著低于无奖惩条件($M = 0.48, SD = 0.31$)，$t(29) = -3.69, p < 0.001$, $t(29) = -3.40, p < 0.001$；社会奖惩与金钱奖惩条件下的欺骗率无显著差异，$t(29) = 0.12, p = 0.909$(图 3)。此外，将性别作为协变量进行协方差分析，结果仍然显示奖惩类型主效应是显著的，$F(2, 28) = 8.37, p = 0.001, \eta_p^2 = 0.374$。上述结果表明，与金钱奖惩相似，社会奖惩也能够有效减少人们的欺骗行为，且两者效果无差异。

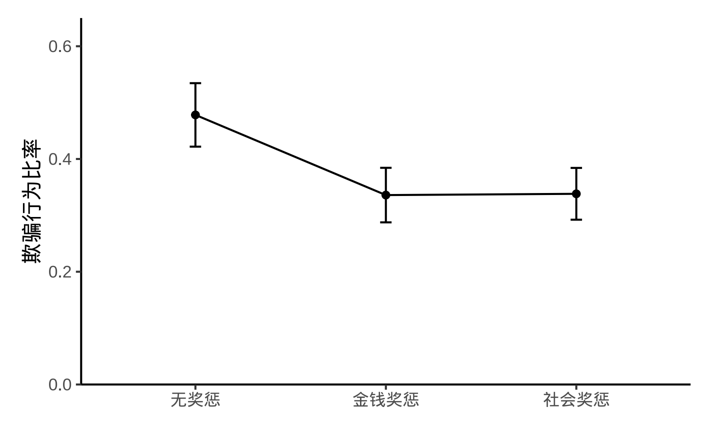


图 3 不同奖惩类型条件下的欺骗率。误差线代表标准误。

为了进一步理解社会奖惩反馈对随后欺骗行为的影响，我们对比了被试在收到社会奖惩反馈后的欺骗比率以及无反馈条件下的欺骗比率。其中，社会奖惩反馈后的欺骗比率是指在社会奖励或社会惩罚条件下，被试在收到相应反馈后的下一个试次中选择欺骗的次数除以该

区块中所有社会奖惩反馈后的试次总数，这并非仅基于一个试次的数据，而是基于所有相关试次的平均值。无反馈条件下的欺骗比率是指在无奖惩反馈条件下被试选择欺骗的次数除以该区块中所有无反馈试次的总数。方差分析结果显示，三种条件下欺骗比例差异显著，$F(2, 48) = 24.50$, $p < 0.001$, $\eta_p^2 = 0.51$。相比于无奖惩条件($M = 0.57$, $SD = 0.24$)，被试在收到社会奖励($M = 0.44$，$SD = 0.24$)或社会惩罚($M = 0.27$, $SD = 0.22$)后的欺骗比率显著降低，$t(24) = −2.47$, $p = 0.021$, $t(24) = −7.29$, $p < 0.001$。此外，被试在收到社会惩罚后的欺骗比率显著低于收到社会奖励后的欺骗比率，$t(24) = −5.16$, $p < 0.001$(图 4)。因此，相比于无社会奖惩反馈，在收到社会奖励或惩罚后，被试在随后的试次中会倾向于选择发送真实的信息，且社会惩罚比社会奖励的干预效果更好。

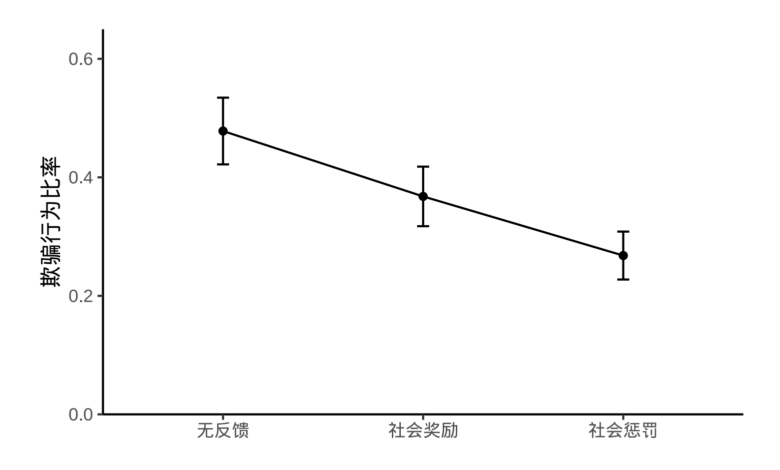


图 4 被试收到社会奖惩反馈后，下一个试次中的欺骗比率与与无反馈条件下的欺骗比率。

此外，对社会奖惩、金钱奖惩、无奖惩三种条件下被试选择诚实和欺骗的反应时进行两因素重复测量方差分析。结果发现，奖惩类型的主效应不显著，$F(2, 44) = 0.51$, $p = 0.604$, $\eta_p^2 = 0.02$；选择类型的主效应不显著，$F(1, 22) = 0.44$, $p = 0.512$, $\eta_p^2 = 0.02$。两者之间的交互作用显著，$F(2, 44) = 5.89$, $p = 0.005$, $\eta_p^2 = 0.21$。简单效应分析表明，在无奖惩条件，选择诚实的反应时($M = 5.31$, $SD = 2.49$)长于选择欺骗的反应时($M = 4.40$, $SD = 2.18$)，$t(24) = 3.43$, $p = 0.002$；但在社会奖惩与金钱奖惩下，选择诚实的反应时与选择欺骗的反应时之间没有显著差异。

漂移扩散模型拟合结果表明，模型参数的痕迹收敛较好，R-hat 值小于 1.05，表明模型拟合良好。模型比较结果表明，四个参数ν, α, z, τ随不同奖惩条件变化的模型最优，留一交叉验证的期望对数预测密度$elpd_loo$ (expected log predictive density for Leave-One-Out cross validation)值最大。模型参数分析结果表明，社会奖惩($M = −0.19$, 95% HDI [−0.23, −0.14])与金钱奖惩($M = −0.18$, 95% HDI [−0.23, −0.14])条件下漂移率ν显著小于无奖惩条件下的漂移

率v(M = −0.03, 95% HDI [−0.07, 0.02])，表明社会奖惩与金钱奖惩促使个体更倾向累积不欺骗的证据。此外，社会奖惩(M = 0.95, 95% HDI [0.89, 1.01])与金钱奖惩(M = 1.03, 95% HDI [0.98, 1.10])条件下非决策时间τ显著长于无奖惩条件下的非决策时间τ(M = 0.75, 95% HDI [0.69, 0.80]) (见图 5)。

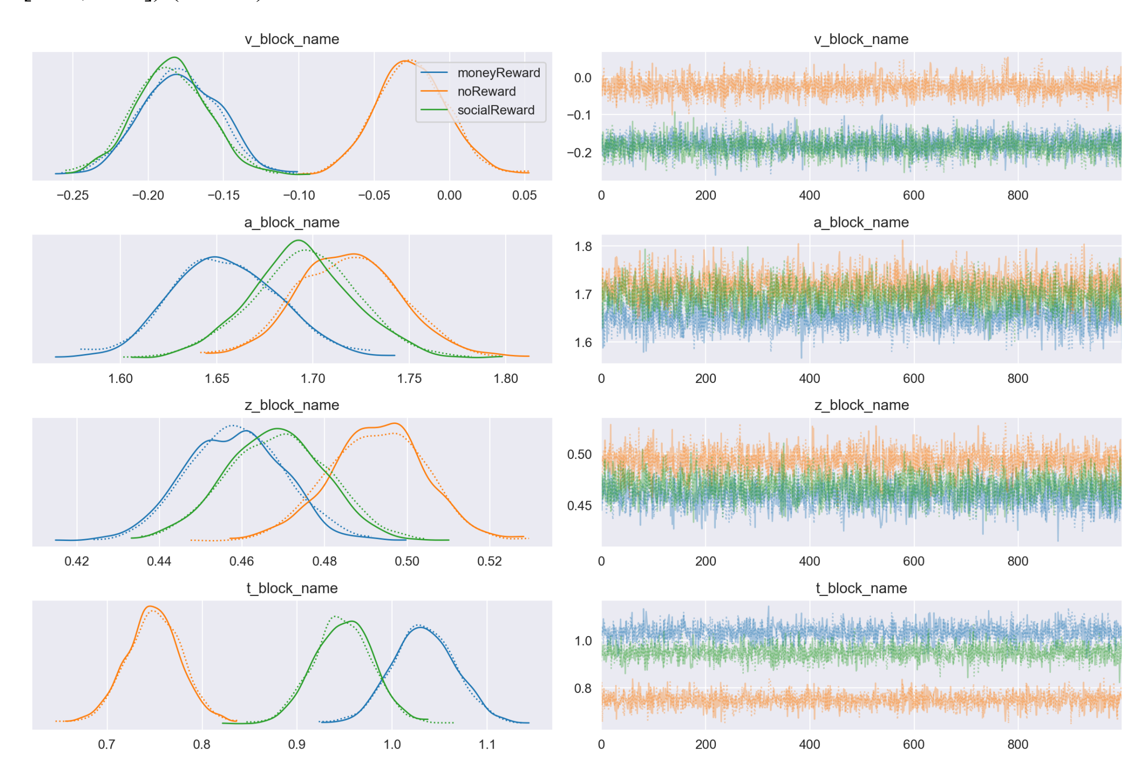


图 5 左边一列为 DDM 的 4 个参数在不同奖惩条件下的后验概率密度分布。从上往下依次为漂移率 v、边界高度 α、起始点偏差 z 以及非决策时间 τ。图中橙色线代表无奖惩条件，绿色线代表社会奖惩条件、蓝色线代表金钱奖惩条件。右边一列为相应模型参数的后验痕迹图(posterior traces plot)。

3 实验 2 社会奖惩对欺骗行为的影响：声誉关注的中介作用

实验 1 发现，与金钱奖惩类似，社会奖惩也能够减少个体的欺骗行为。那么，社会奖惩可能通过什么机制影响人们的欺骗行为呢？社会奖惩涉及社会评价过程，这可能会使个体更加关注自己的声誉从而影响其欺骗行为。为此，实验 2 拟探讨声誉关注在社会奖惩影响欺骗行为的过程中是否起到中介作用。

3.1 被试

为了考察中介效应，招募了 60 名在校大学生，平均年龄为 20.25(SD = 1.74)，其中男生 23 人，女生 37 人。

3.2 实验设计

本实验为单因素的被试内实验设计，自变量为奖惩类型(社会奖惩、金钱奖惩、无奖惩)，

因变量为被试欺骗行为的比例，中介变量为声誉关注。

3.3 实验材料与任务

我们改编了 Wu, Balliet 和 Lange(2016)研究中使用的声誉关注问卷，以测量被试对自己的声誉关注程度。该问卷由六个题目构成，具有良好的信效度。根据该研究社会奖惩的性质从中选取了两个题目来测量被试的声誉关注程度，采用 9 级评分*(1=完全不同意，9=完全同意)*，其中第二题为反向计分，评分越高说明被试对自己的声誉越关注(附录 1)。

3.4 实验程序

实验程序与实验 1 基本相同，被试完成信号博弈任务，在每一个区块结束时进行声誉关注的测量。

3.5 结果

首先，对社会奖惩、金钱奖惩、无奖惩三种条件下被试的情绪水平进行重复测量方差分析。结果发现，三种条件下的被试的情绪无显著差异，$F(2, 118) = 2.05, p = 0.133, \eta_p^2 = 0.03$。其次，对社会奖惩、金钱奖惩、无奖惩三种条件下被试的欺骗率进行重复测量方差分析。结果发现，被试在三种条件下的欺骗率存在显著差异，$F(2, 118) = 12.99, p < 0.001, \eta_p^2 = 0.18$。其中，社会奖惩($M = 0.35, SD = 0.25$)和金钱奖惩($M = 0.33, SD = 0.25$)条件下的欺骗率显著低于无奖惩条件($M = 0.46$，$SD= 0.30$), $t(59) = −3.87, p < 0.001$, $t(59) = −4.33, p < 0.001$；社会奖惩与金钱奖惩条件下的欺骗率无显著差异，$t(59) = 1.01, p = 0.315$(图 6)。此外，将性别作为协变量进行协方差分析，结果仍然显示奖惩类型主效应是显著的，$F(2, 116) = 11.45, p < 0.001$, $\eta_p^2 = 0.17$。结果表明，与金钱奖惩相似，社会奖惩也能够减少人们的欺骗行为，且两者效果无显著差异。

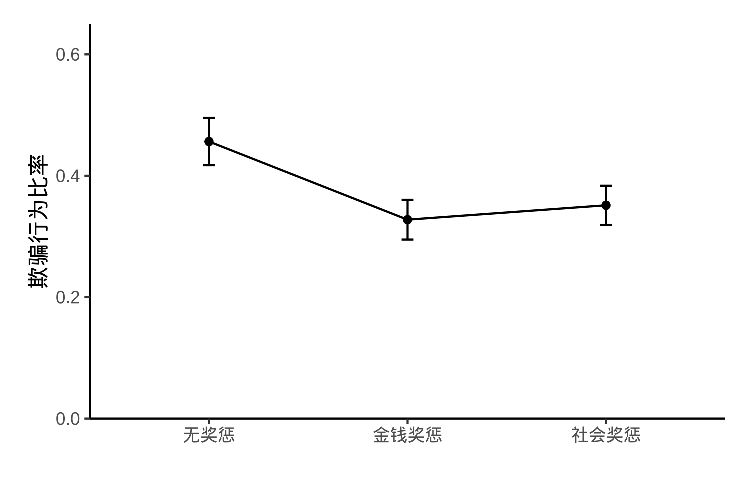


图 6 不同类型的奖惩对欺骗行为的影响

为分析声誉关注在社会奖惩对欺骗行为影响的过程中的中介效应，以有无社会奖惩(社

会奖惩= 1，无奖惩= 0)作为自变量，以声誉关注得分作为中介变量，欺骗率作为因变量。采用 PROCESS 3.1 检验“有无社会奖惩→声誉关注→欺骗行为”这一中介路径，选择 Model 4，抽样量设为 5000。结果发现，表明该路径的中介效应显著，Mean bootstrapped indirect effect *ab* = –0.12, *BootSE* = 0.04，95%置信区间为[*LLCI* = −0.22, *ULCI* = −0.05]，不包含 0(见图 7)。

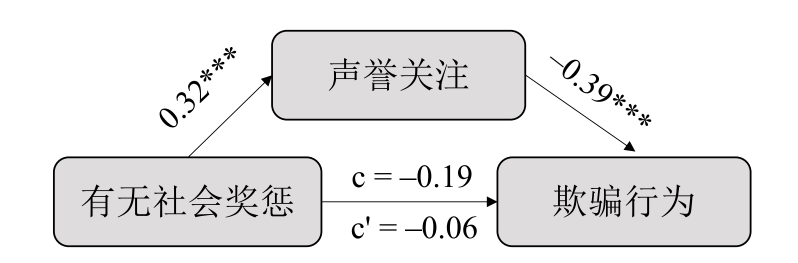


图 7 社会奖惩中介效应路径图

注: *表示 $p < 0.05$, **表示 $p < 0.01$, ***表示 $p < 0.001$, 下同。

此外，我们也分析了声誉关注在金钱奖惩对欺骗行为影响过程中的中介效应。以有无金钱奖惩(金钱奖惩= 1，无奖惩= 0)作为自变量，以声誉关注得分作为中介变量，欺骗率作为因变量，检验“有无金钱奖惩→声誉关注→欺骗行为”这一中介路径。结果发现，该路径 95%置信区间为[*LLCI* = −0.10, *ULCI* = 0.02]，置信区间包含 0，中介效应不显著，如图 8 所示。

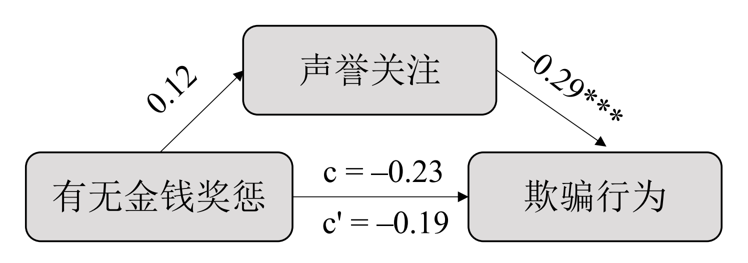


图 8 金钱奖惩中介效应路径图

此外，对社会奖惩、金钱奖惩、无奖惩三种条件下被试选择诚实和欺骗的反应时进行两因素重复测量方差分析。结果发现，奖惩类型的主效应不显著，$F(2, 94) = 0.12, p = 0.892, \eta_p^2 = 0.002$；选择类型的主效应不显著，$F(1, 47) = 0.07, p = 0.788, \eta_p^2 = 0.002$。两者之间的交互作用显著，$F(2, 94) = 6.72, p = 0.002, \eta_p^2 = 0.13$。简单效应分析表明，在无奖惩条件，选择诚实的反应时($M = 4.69, SD = 2.51$)长于选择欺骗的反应时($M = 4.21, SD = 2.28$)，$t(47) = 2.98, p = 0.005$。在金钱奖惩下，选择诚实的反应时($M = 4.35, SD = 1.69$)短于选择欺骗的反应时($M = 4.78, SD = 2.41$)，$t(47) = 2.04, p = 0.047$。在社会奖惩下，选择诚实的反应时与选择欺骗的反应时之间没有显著差异，$t(47) = 0.75, p = 0.459$。

漂移扩散模型拟合结果表明，模型参数的痕迹收敛较好，R-hat 值小于 1.05，表明模型

拟合良好。模型比较结果表明，四个参数v, α, z, τ随不同奖惩条件变化的模型最优，$elpd_loo$值最大。模型参数分析结果表明，社会奖惩($M = -0.16$, 95% HDI [−0.19, −0.13])与金钱奖惩($M = -0.21$, 95% HDI [−0.24, −0.17])条件下漂移率v显著小于无奖惩条件下的漂移率v($M = -0.05$, 95% HDI [−0.08, 0.02])，表明社会奖惩与金钱奖惩促使个体更倾向累积不欺骗的证据。此外，社会奖惩($M = 0.92$, 95% HDI [0.86, 0.97])与金钱奖惩($M = 0.94$, 95% HDI [0.88, 0.99])条件下非决策时间τ显著长于无奖惩条件下的非决策时间τ($M = 0.76$, 95% HDI [0.71, 0.80])(见图 9)。

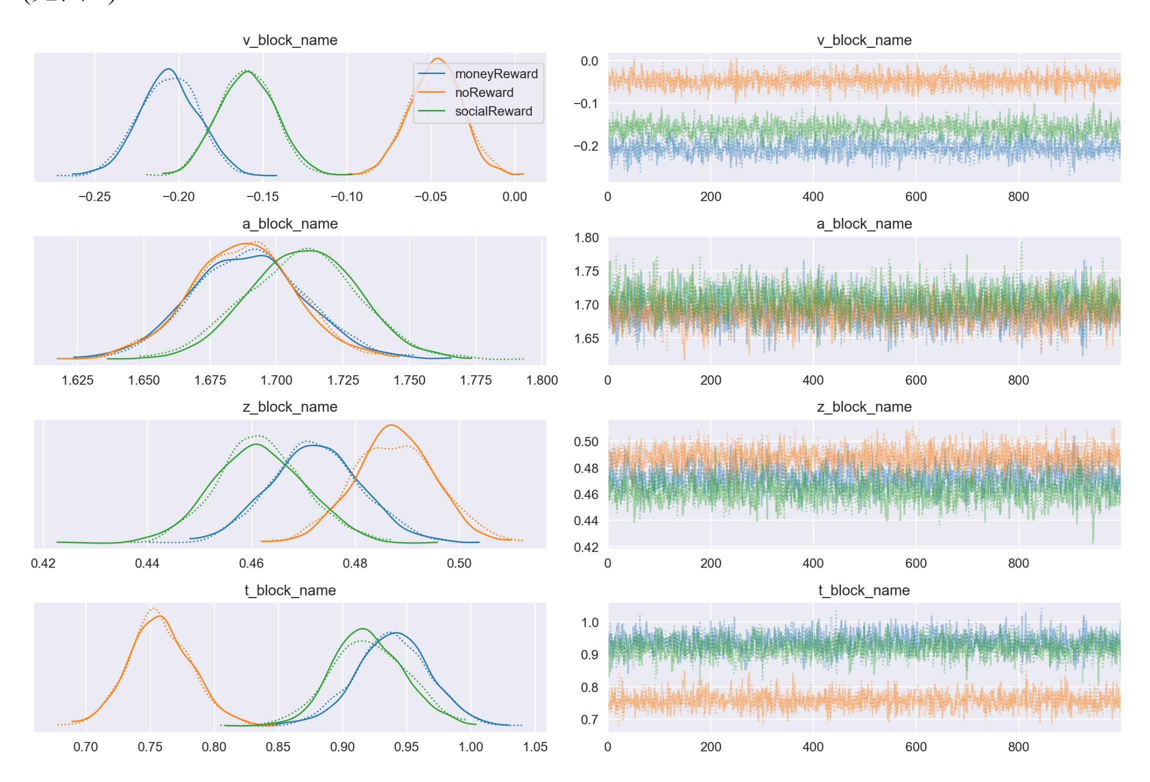


图 9 左边一列为 DDM 的 4 个参数在不同奖惩条件下的后验概率密度分布。从上往下依次为漂移率 v、边界高度 α、起始点偏差 z 以及非决策时间 τ。图中橙色线代表无奖惩条件，绿色线代表社会奖惩条件、蓝色线代表金钱奖惩条件。右边一列为相应模型参数的后验痕迹图(posterior traces plot)。

4 实验 3 社会奖惩对欺骗行为的影响：社会价值取向的调节作用

实验 2 发现社会奖惩能够增加个体对自身的声誉关注，从而减少其欺骗行为，而金钱奖惩则不能。社会奖惩影响个体行为的过程可能会受到个体人格特质的调节，其中社会价值取向是重要的因素之一。相比于亲自我取向个体，亲社会取向个体表现出对社会奖惩(接纳或拒绝)更强的敏感性。因此，实验 3 拟探讨社会价值取向是否会调节社会奖惩通过声誉关注影响欺骗行为的中介过程。

4.1 被试

共 193 名被试参与了本实验，采用社会价值取向滑块测验(the SVO slider measure)测量其社会价值取向，其中 125 名为亲自我者(proself)，68 名为亲社会者(prosocial)，平均年龄为 22.83(SD = 3.75)，男生 96 人，女生 97 人。

4.2 实验设计

采用 2(社会价值取向：亲自我取向，亲社会取向)×3(奖惩类型：社会奖惩，金钱奖惩，无奖惩)两因素混合实验设计，其中奖惩类型为被试内变量，社会价值取向为被试间变量，因变量为被试欺骗行为的比例。

4.3 实验材料与任务

(1)社会价值取向测量

采用由 Murphy，Ackermann 和 Handgraaf(2011)编制社会价值取向滑块测验测量个体的社会价值取向，该测验共 15 个题目，分为 6 个初级项目和 9 个次级项目。每个题目都给出对 9 个金额(如 150 元)分配选项，决策者需从中选择自己偏好的分配选项(见附录 B)。初级项目是滑块测验的主体，用于评定个体的 SVO 角度或类型，即利他取向、亲社会取向、个人取向和竞争取向，其角度值越大表示决策者越在意他人的收益，最大值 61.39°代表纯粹的利他取向，最小值–16.26°代表纯粹的竞争取向(张振 等, 2015)。次级项目是可选项目，用于检测个体的不平等厌恶动机。本实验采用最常用的两种社会价值取向类型：(1)亲自我取向，包括个人倾向和竞争倾向；(2)亲社会取向，包括利他和亲社会取向。

(2)声誉关注测量

同实验 2。

4.4 实验程序

首先，通过社会价值取向滑块测验测量被试的社会价值取向 SVO 角度。然后，根据 SVO 角度筛选亲自我取向和亲社会取向的被试进行后续实验。其中，SVO 角度>22.45°为亲社会取向者，SVO 角度<22.45°为亲自我取向者(Murphy et al., 2011)。随后的实验程序与实验 2 相同，采用信号博弈任务，并在每一个区块结束后测量被试的声誉关注。

4.5 结果

对奖惩类型与不同社会价值取向被试的欺骗比例进行两因素方差分析。结果发现，奖惩类型的主效应显著，$F(2, 382) = 40.66, p < 0.001, \eta_p^2 = 0.183$。社会价值取向的主效应显著，$F(1, 191) = 30.99, p < 0.001, \eta_p^2 = 0.14$，亲社会取向被试的欺骗比例($M = 0.35, SD = 0.29$)显著低于亲自我取向被试的欺骗比例($M = 0.57, SD = 0.33$)。社会价值取向与奖惩类型的交互作用显著，$F(2, 382) = 5.44, p = 0.005, \eta_p^2 = 0.03$。简单效应分析表明，对于亲自我取向个体，社

会奖惩的主效应显著，$F(2, 191) = 19.66, p < 0.001, \eta_p^2 = 0.17$。多重比较发现，社会奖惩条件下的欺骗比例($M = 0.60, SD = 0.32$)与无奖惩条件($M = 0.65, SD = 0.31$)的欺骗比例没有显著差异，$t(191) = -1.76, p = 0.079$；金钱奖惩条件下的欺骗比例($M = 0.46, SD = 0.34$)显著低于无奖惩条件($M = 0.65, SD = 0.31$)，$t(191) = -5.98, p < 0.001$。而对于亲社会取向个体，社会奖惩的主效应显著，$F(2, 191) = 23.70, p = < 0.001, \eta_p^2 = 0.20$。社会奖惩条件下的欺骗比例($M = 0.32, SD = 0.27$)显著低于无奖惩条件($M = 0.45, SD = 0.32$)，$t(191) = -5.95, p < 0.001$；金钱奖惩条件下的欺骗比例($M = 0.30, SD = 0.26$)显著低于无奖惩条件($M = 0.45, SD = 0.32$)，$t(191) = -6.19, p < 0.001$；社会奖惩与金钱奖惩的差异条件下被试的欺骗比例无显著差异，$t(191) = 0.81, p = 0.420$(见图 10)。

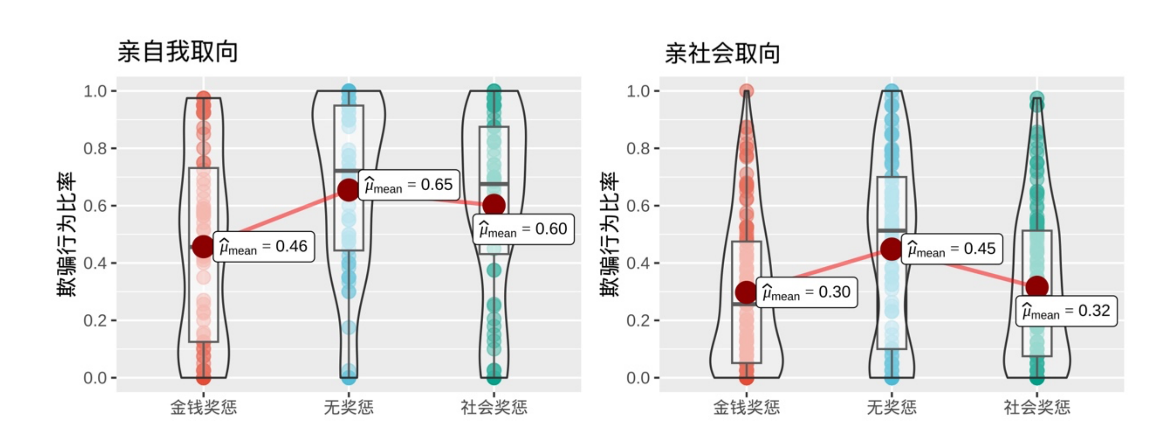


图 10 亲自我取向与亲社会取向个体在不同奖惩类型条件下的欺骗率。图中红色的实心圆点代表该条件下的均值，每个半透明的圆点代表每个被试的欺骗率，两侧对称的形状表示欺骗率的概率分布。

首先，再次分析了声誉关注在社会奖惩对欺骗行为影响的过程中的中介效应，以有无社会奖惩(社会奖惩= 1，无奖惩= 0)作为自变量，以声誉关注得分作为中介变量，欺骗率作为因变量，检验“有无社会奖惩→声誉关注→欺骗行为”这一中介路径。结果发现，表明该路径的中介效应显著，95%置信区间为[$LLCI = -0.092, ULCI = -0.032$]，不包含 0。此外，我们也分析了声誉关注在金钱奖惩对欺骗行为影响过程中的中介效应。结果发现，表明该路径的中介效应不显著，95%置信区间为[$LLCI = -0.012, ULCI = 0.031$]，包含 0。

此外，采用 PROCESS 中的 Model 8 检验社会价值取向是否调节声誉关注在社会奖惩影响欺骗行为中的中介作用(有调节的中介模型)。结果发现，社会奖惩正向预测声誉关注($\beta = 0.88, p < 0.001$)，负向预测欺骗行为($\beta = -0.05, p < 0.001$)；社会价值取向正向预测声誉关注($\beta = 0.82, p < 0.001$)，负向预测欺骗行为($\beta = -0.20, p < 0.001$)。社会奖惩和社会价值取向的交互项对声誉关注有显著的预测作用($\beta = 0.90, p = 0.030$)(见表 1)。具体而言，对于亲自我

取向的个体，声誉关注在社会奖惩与欺骗行为间的中介效应为 ab = –0.018, $BootSE$ = 0.021，z = –0.82, p = 0.415, 95%置信区间为 [$LLCI$ = −0.061, $ULCI$ = 0.022]，中介效应不显著；而对于亲社会取向的个体，声誉关注在社会奖惩与欺骗行为间的中介效应为 ab = –0.070, $BootSE$ = 0.016，z = –4.29, p < 0.001, 95%置信区间为 [$LLCI$ = −0.104, $ULCI$ = −0.040]，中介效应显著。

表 1 社会价值取向的调节效应分析结果

结果变量	预测变量	整体拟合指标		回归系数显著性	
		R^2	***F***	**β**	***t***
声誉关注		0.10	13.64***		
	社会奖惩			0.88	4.50***
	社会价值取向			0.82	3.99***
	社会奖惩*社会价值取向			0.90	2.18*
欺骗行为		0.27	35.82***		
	声誉关注			-0.06	-7.79***
	社会奖惩			-0.05	-1.83
	社会价值取向			-0.20	-6.48***
	社会奖惩*社会价值取向			-0.03	-0.45

漂移扩散模型拟合结果表明，模型参数的痕迹收敛较好，R-hat 值小于 1.05，表明模型拟合良好。模型比较结果表明，四个参数ν, α, z, τ随不同奖惩条件变化的模型最优，$elpd_loo$值最大。模型参数分析结果表明，社会奖惩(M = −0.13, 95% HDI [−0.14, −0.11])与金钱奖惩(M = −0.22, 95% HDI [−0.24, −0.20])条件下漂移率ν显著小于无奖惩条件下的漂移率ν(M = 0.05, 95% HDI [0.03, 0.06])，表明社会奖惩与金钱奖惩促使个体更倾向累积不欺骗的证据。此外，社会奖惩(M = 0.81, 95% HDI [0.79, 0.83])与金钱奖惩(M = 0.71, 95% HDI [0.70, 0.73])条件下非决策时间τ显著长于无奖惩条件下的非决策时间τ(M = 0.76, 95% HDI [0.71, 0.80]) (见图 12)。

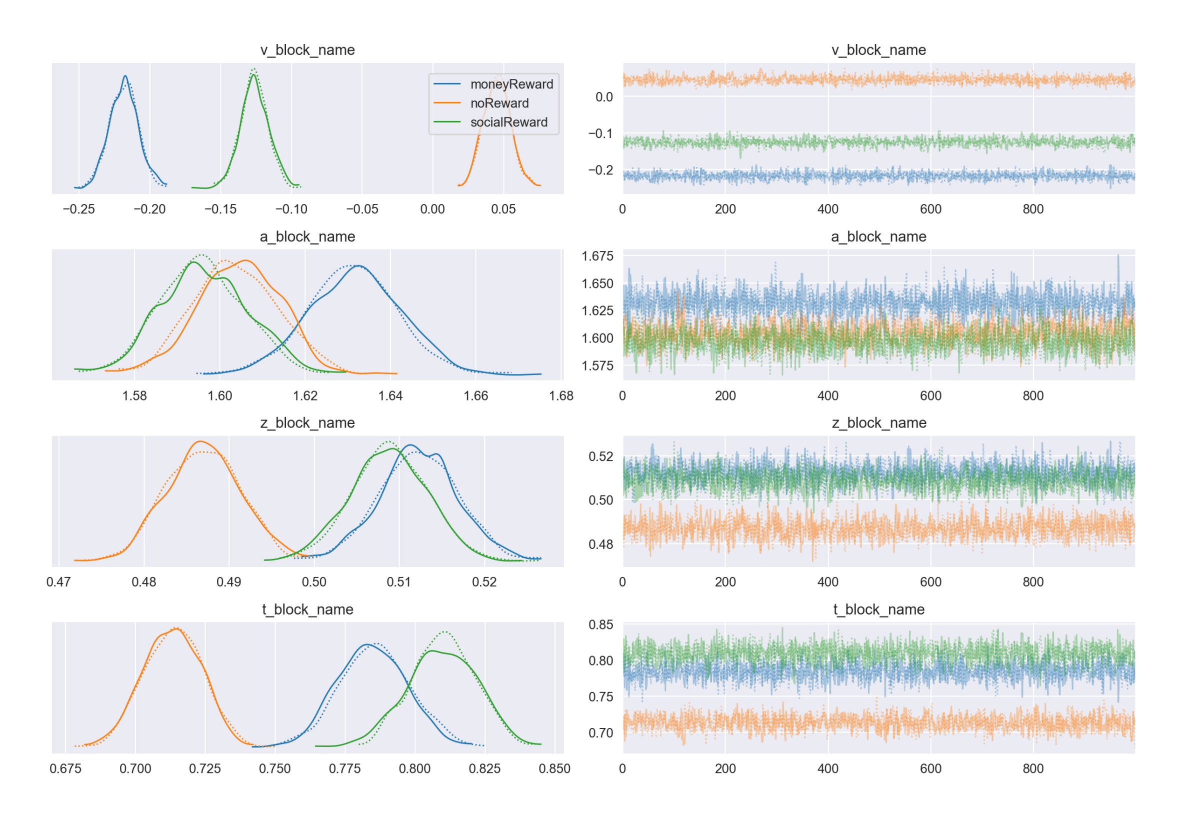


图 12 左边一列为 DDM 的 4 个参数在不同奖惩条件下的后验概率密度分布。从上往下依次为漂移率 v、边界高度 α、起始点偏差 z 以及非决策时间 τ。图中橙色线代表无奖惩条件，绿色线代表社会奖惩条件、蓝色线代表金钱奖惩条件。右边一列为相应模型参数的后验痕迹图(posterior traces plot)。

5 讨论

本研究通过 3 项实验，采用信号博弈任务，系统地探讨了社会奖惩对欺骗行为的影响及其中介与调节机制。结果发现，与金钱奖惩类似，社会奖惩能够减少欺骗行为，且社会惩罚比社会奖励的效果更好。声誉关注在社会奖惩对欺骗行为的影响中起中介作用，社会价值取向调节社会奖惩通过声誉关注影响欺骗行为的中介过程。

5.1 社会奖惩对欺骗行为的影响

本研究发现，社会奖惩与金钱奖惩在减少欺骗行为方面具有相似的效果，均能显著降低个体的欺骗行为。该结果与已有研究一致，即社会奖惩能够增强个体的亲社会行为。当个体遵循社会规范并表现出亲社会行为时，往往会获得来自他人的微笑、表扬、认可、接纳、支持和关爱等社会肯定，这一过程满足了个体的人际关系需求与归属感，进而强化其亲社会行为(吕房艳 等, 2021; Kringelbach & Rolls, 2003)。研究表明，预期他人的口头赞赏可以增加个体的亲社会行为(Ellingsen & Johannesson, 2008)，甚至有人愿意放弃金钱收益以获得社会认同与接纳(Shore & Heerey, 2011)。相反，社会惩罚会引发个体的负面情绪，并激活类似于躯

体伤害的疼痛反应神经系统(Beston, 2019)。为了避免社会排斥、声誉损害等惩罚，个体会表现出更多的亲社会行为。

进一步分析表明，接受社会奖励或社会惩罚后，个体在随后的试次中更倾向于选择发送真实的信息，但相比社会奖励，社会惩罚在抑制欺骗行为方面具有更强的效果。这可能源于社会惩罚激活的情绪反应更为强烈，进而更有效地抑制了不符合社会规范的行为。这一研究结果支持了负性偏差效应(negativity bias effect)，即个体对负面信息的敏感度高于正面信息。研究表明，负性事件在情绪、行为和认知层面产生的影响通常更为持久和强烈，且负性刺激(如，社会惩罚)能够迅速且显著激活个体的情绪反应(Baumeister et al., 2001; Rozin & Royzman, 2001)。相比于社会奖励，社会惩罚更容易引发羞耻、内疚或焦虑等负面情绪，并对个体的心理状态产生更持久的影响(Eisenberger et al., 2003; Kujawa et al., 2015)。

此外，本研究还发现，奖惩类型和选择类型反应时的交互作用显著。在无奖惩条件，选择诚实的反应时长于选择欺骗的反应时；但在社会奖惩与金钱奖惩下，选择诚实的反应时与选择欺骗的反应时之间没有显著差异。Debey 等人(2015)通过分析反应时揭示了诚实和欺骗行为的认知机制，发现诚实行为通常伴随着更长的反应时间，而欺骗行为则通常需要较短的反应时间。这表明，诚实通常涉及较慢的反思性系统，需要仔细考虑道德与社会规范，而欺骗可能更多依赖快速的直觉性系统。在无奖惩条件下，反思性系统更占主导，导致诚实反应时间更长(Greene et al., 2008)。然而，在社会奖惩与金钱奖惩条件下，明确的动机(获得奖励或避免惩罚)简化了决策过程，减少了反思性系统的介入(Gino & Ariely, 2012)，促使个体更快做出诚实选择。漂移扩散模型的结果进一步支持了上述认知过程，在社会奖惩与金钱奖惩条件下，漂移率v显著低于无奖惩条件，且主要集中在负值区间。漂移率v指信息积累的速度，反映了个体在决策过程中对某一选项的偏好程度(袁博 等, 2023)，越负的漂移率表明个体表明个体越重视诚实选项。因此，社会奖惩与金钱奖惩都促使个体更倾向累积诚实的证据，进而做出诚实选择。

5.2 声誉关注在社会奖惩影响欺骗行为中的中介作用

本研究发现，声誉关注在社会奖惩影响欺骗行为中起到中介作用。具体而言，社会奖惩能够增加个体的声誉关注，从而减少其欺骗行为；而金钱奖惩则无法通过声誉关注影响欺骗行为。这一结果支持了间接互惠理论(indirect reciprocity theory)和利他声誉理论(altruistic reputation theory)。间接互惠理论认为，声誉是解释人类利他主义的关键概念(Nowak & Sigmund, 2005)，在声誉系统下，个体会关注自身的声誉和他人对自己的印象，声誉成为评估行为的标准之一(Leimar & Hammerstein, 2001)。利他声誉理论认为，人们通过公开展示利

他行为(如帮助他人、慷慨捐赠等)建立积极的声誉，而这种声誉能带来社交和生存上的优势(Trivers, 1971)。相比于金钱奖惩，社会奖惩涉及更多的社会评价过程，个体更加关注自己的声誉，从而影响其欺骗行为。声誉关注是指由他人对自己行为的评估所激活的关注目标，促使人们调整行为以获得良好声誉，从而增加未来获取长远收益的可能性(Sperber & Baumard, 2012)。已有研究表明，当个体能够通过间接互惠获得声誉激励时，他们更愿意合作(Milinski et al., 2006)。此外，还有研究发现，个体会根据群体其他成员的评价调整自己的合作行为，以便被群体或组织接受和认可(Brady et al., 2017; Sommerfeld et al., 2007)。当人们预期自己的声誉信息被传播时，他们会更关注他人意见，促使自己做出更多的贡献；而当接收到其他群体成员对自己的负面评价时，他们会意识到这些评价会对声誉产生不利影响，从而增加对团队的贡献并遵守团体规范(Wu et al., 2016)。

5.3 社会价值取向调节社会奖惩对欺骗行为的影响

本研究还发现，社会价值取向调节了社会奖惩对欺骗行为的影响。对于亲自我取向的个体，金钱奖惩能够显著降低欺骗比率，但社会奖惩在减少欺骗行为方面效果有限。相反，对于亲社会取向的个体，社会奖惩与金钱奖惩均能显著减少其欺骗行为。此外，社会价值取向调节了社会奖惩通过声誉关注影响欺骗行为的中介过程。以往研究表明，亲自我取向者更加重视自身利益，而亲社会取向者更加重视群体利益，在面临社会困境时会表现出更多的合作等亲社会行为(刘长江, 郝芳, 2011)。社会价值取向通过影响个体的声誉关注，进而调节他们在面对奖惩时的欺骗行为。声誉被视为亲社会取向个体维持社会联系的重要机制，对于亲社会个体而言，声誉不仅是一种外在的认可，更是维持社交关系和促进合作的一种内在驱动力(De Cremer & Van Lange, 2001)。亲社会取向的个体更关注社会和他人期望，倾向于在人际互动情境中表现出更高的声誉关注(Van Lange, 1999; Simpson & Willer, 2008; Cameron & Payne, 2011)。声誉管理理论认为，个体会通过操控自己在他人眼中的形象来维护其社会地位。社会价值取向影响个体对声誉损益的敏感度，相比于亲自我取向的个体，亲社会取向的个体更可能通过社会奖惩来调节其行为，以避免因不诚实行为而导致的声誉损失(Leary & Kowalski, 1990)。这表明亲社会取向者在社会奖惩情境中可能更遵守社会规范，以维护自身声誉，减少欺骗行为；而亲自我取向者则可能对社会奖惩的声誉威胁反应较弱，表现出较高的欺骗倾向。

5.4 理论贡献与实践启示

本研究揭示了社会奖惩相较于金钱奖惩在抑制欺骗行为方面的独特机制，尽管两者均能有效减少欺骗行为，但其影响路径存在显著差异。金钱奖惩主要通过提供外部激励或增加潜

在成本减少欺骗行为(Gneezy, 2005)；而社会奖惩则依赖个体对声誉的关注，激发内在的道德责任感，从而抑制欺骗行为(Feinberg et al., 2014)。虽然金钱奖惩能通过物质激励影响个体决策，但其作用往往受外部条件限制，而社会奖惩基于个体对声誉的关注，这种内在驱动力具有更强的持久性(Feinberg et al., 2014)。本研究发现，即便在使用表情符号作为社会奖惩手段的情境下，个体仍表现出对声誉的关注，并减少欺骗行为。这一发现进一步证实了社会奖惩的广泛适用性，表明即便是轻量的社会信号也能影响道德决策。这为未来关于社会行为的研究提供了新的视角，并突显了社会奖惩机制在道德规范维持中的重要性。

本研究为如何有效抑制欺骗行为提供了重要启示，研究结果表明，社会奖惩作为一种更具潜力且低成本的手段，也能有效抑制欺骗行为。在学校教育背景下，教育工作者应重视社会奖惩在德育中的作用。传统的教育管理往往依赖物质奖励来激励学生诚实守信。但本研究表明，社会奖惩也能有效减少不道德行为。由于青少年的价值观尚处于发展阶段，过度依赖金钱奖励可能导致短期功利取向，而忽视内在道德责任感的培养(Frey & Jegen, 2001)。因此，社会奖惩能对学生的价值观的塑造产生更深远的影响。研究表明，金钱奖励可能削弱个体的内在动机，使行为更加依赖外部激励，而非道德信念的驱动(Deci et al., 1999)。相比之下，社会奖励能满足个体的归属感和自我认同需求(Ryan & Deci, 2000)，增强学生对诚信与社会规范的内在认同，从而促进更持久的道德行为(Feinberg et al., 2014)。因此，在教育管理中，适当强调社会奖惩不仅有助于塑造学生的道德价值观，还能减少其对物质奖励的过度依赖，使诚信行为更具可持续性。

5.5 研究局限与未来展望

本研究还存在一些不足之处，这也为未来研究提供了发展方向。首先，社会奖惩包括多种形式，如尊重、认可、称赞、接纳，或批评、反对、排斥、散布流言等(Ramirez-Marin & Shafa, 2018; Kim & Jeong, 2020)。然而，当前关于社会奖惩的具体范畴仍不明确，且缺乏系统性的结构分析。本研究仅采用了微笑和沮丧表情作为社会奖惩的操作方式，其他形式的社会奖惩对欺骗行为的影响仍需进一步探讨，以全面理解社会奖惩的作用机制。其次，本研究仅关注了社会价值取向这一个体变量，而其他个体差异也可能会调节社会奖惩对欺骗行为的影响。例如，研究发现，社会奖惩敏感性较高的个体更关注社会评价和声誉维护(Gino & Pierce, 2009)，而抑郁人群与健康人群的社会奖赏加工也存在差异(李思瑾 等, 2024)，未来研究可以进一步探讨这些个体差异的调节作用，以更全面地理解个体特征如何影响社会奖惩的效果。最后，欺骗行为通常发生在特定的社会情境中(Gneezy, 2005; Mazar et al., 2008)，未来研究可进一步考察不同社会情境对社会奖惩效果的调节作用。例如，社会排斥可能通过威胁个体的

归属感，进而影响其对社会奖惩的反应(Eisenberger et al., 2003)。此外，情境公开性也可能会影响社会奖惩的效果。在公开的情境中，个体的自我意识更强，与自我相关的神经区域更活跃(Somerville et al., 2013)，使得个体更容易感受到被评价，从而更加关注声誉。因此，未来研究可进一步探讨社会情境因素如何塑造社会奖惩的影响模式。

6 结论

相对于无奖惩条件，社会奖惩与金钱奖惩均能减少被试的欺骗行为，且社会惩罚比社会奖励的效果更好；社会奖惩能够增加个体对自身的声誉关注，从而减少其欺骗行为。个体的社会价值取向调节了社会奖惩通过声誉关注影响欺骗行为这一中介过程，社会奖惩使亲社会取向个体更加关注自己的声誉从而减少其欺骗行为。

参考文献

Allingham, M. G., & Sandmo, A. (1972). Income tax evasion: A theoretical analysis. *Journal of Public Economics*, *1*(3–4), 323–338.

Ariely, D., & Gino, F. (2012). Cheating, self-signaling, and groups: Reminding people of their moral standards. *Social Psychological and Personality Science*, *3*(3), 344–352.

Balliet, D., Parks, C., & Joireman, J. (2009). Social value orientation and cooperation in social dilemmas: A meta-analysis. *Group Processes & Intergroup Relations*, *12*(4), 533–547.

Balliet, D., & Van Lange, P. A. (2013). Trust, conflict, and cooperation: a meta-analysis. *Psychological Bulletin*, *139*(5), 1090–1112.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, *5*(4), 323–370.

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, *76*(2), 169–217.

Behnk, S., Barreda-Tarrazona, I., & Garcia-Gallego, A. (2018). Punishing liars–How monitoring affects honesty and trust. *PLoS One*, *13*(10), e0205420.

Berman, J. Z., Levine, E. E., Barasch, A., & Small, D. A. (2015). The Braggart's dilemma: On the social rewards and penalties of advertising prosocial behavior. *Journal of Marketing Research*, *52*(1), 90–104.

Beston, P. (2019). *The effect of social rewards and punishments on learning and cooperative decision-making* (Master's thesis). Bangor University, United Kingdom.

Boutet, I., LeBlanc, M., Chamberland, J. A., & Collin, C. A. (2021). Emojis influence emotional communication, social attributions, and information processing. *Computers in Human Behavior*, *119*, 106722.

Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., & Van Bavel, J. J. (2017). Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences*, *114*(28), 7313–7318.

Cameron, C. D., & Payne, B. K. (2011). Escaping affect: how motivated emotion regulation creates insensitivity to mass suffering. *Journal of Personality and Social Psychology*, *100*(1), 1–15.

Carr, Z. M., Solbu, A., & Frank, M. G. (2019). Why methods matter: Approaches to the study of deception and considerations for the future. In T. Docan-Morgan (Ed.), *The Palgrave Handbook of Deceptive Communication* (pp. 267–286). Palgrave Macmillan.

Cherbonnier, A., & Michinov, N. (2021). The recognition of emotions beyond facial expressions: Comparing emoticons specifically designed to convey basic emotions with other modes of expression. *Computers in*

Human Behavior, *118*, 106689.

Cinyabuguma, M., Page, T., & Putterman, L. (2005). Cooperation under the threat of expulsion in a public goods experiment. *Journal of Public Economics*, *89*(8), 1421–1435.

De Cremer, D., & Van Lange, P. A. (2001). Why prosocials exhibit greater cooperation than proselfs: The roles of social responsibility and reciprocity. *European Journal of Personality*, *15*(S1), S5–S18.

De Cremer, D., & Tyler, T. R. (2005). Managing group behavior: The interplay between procedural justice, sense of self, and cooperation. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology*, *37*(5), 151–218.

Debey, E., Ridderinkhof, R. K., De Houwer, J., De Schryver, M., & Verschuere, B. (2015). Suppressing the truth as a mechanism of deception: Delta plots reveal the role of response inhibition in lying. *Consciousness and Cognition*, *37*, 148–159.

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology*, *18*(1), 105–115.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, *125*(6), 627–668.

Depaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, *129*(1), 74–118.

Eisenberger, N. I., Lieberman, M. D., & Williams, K. D. (2003). Does rejection hurt? An fMRI study of social exclusion. *Science*, *302*(5643), 290–292.

Ellingsen, T., & Johannesson, M. (2008). Pride and Prejudice: The human side of incentive theory. *American Economic Review*, *98*(3), 990–1008.

Feinberg, M., Willer, R., & Schultz, M. (2014). Gossip and ostracism promote cooperation in groups. *Psychological Science*, *25*(3), 656–664.

Fischer, B., & Herbert, C. (2021). Emoji as affective symbols: affective judgments of emoji, emoticons, and human faces varying in emotional content. *Frontiers in Psychology*, *12*, 645173.

Flynn, F. J., Reagans, R. E., Amanatullah, E. T., & Ames, D. R. (2006). Helping one's way to the top: Self-monitors achieve status by helping others and knowing who helps whom. *Journal of Personality and Social Psychology*, *91*(6), 1123–1137.

Frey, B. S., & Jegen, R. (2001). Motivation crowding theory. *Journal of Economic Surveys*, *15*(5), 589–611.

Gächter, S., & Herrmann, B. (2009). Reciprocity, culture and human cooperation: previous insights and a new cross-cultural experiment. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1518)*,* 791–

806.

Gino, F., & Pierce, L. (2009). Dishonesty in the name of equity. *Psychological Science*, *20*(9), 1153–1160.

Gintis, H. (2000). Strong reciprocity and human sociality. *Journal of Theoretical Biology*, *206*(2), 169–179.

Gneezy, U. (2005). Deception: The role of consequences. *American Economic Review*, *95*(1), 384–394.

Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition*, *107*(3), 1144–1154.

Grosch, K., & Rau, H. A. (2017). Gender differences in honesty: The role of social value orientation. *Journal of Economic Psychology*, *62*, 258–267.

Guala, F. (2012). Reciprocity: weak or strong? What punishment experiments do (and do not) demonstrate. *Behavioral and Brain Sciences*, *35*(1), 1–15.

Hand, C. J., Kennedy, A., Filik, R., Pitchford, M., & Robus, C. M. (2023). Emoji identification and emoji effects on sentence emotionality in ASD-diagnosed adults and neurot-ypical controls. *Journal of Autism and Developmental Disorders*, *53*(6), 2514–2528.

Hardy, C. L., & Van Vugt, M. (2006), Nice guys finish first: The competitive altruism hypothesis. *Personality and Social Psychology Bulletin*, *32*(10), 1402–1413.

Jin, Y. C., Deng, C. L., Wu, P., Lin, X., Zheng, P. X., & An, J. X. (2022). Emoji image symbol's social function and application. *Advances in Psychological Science*, *30*(5), 1062–1077.

[靳宇倡, 邓成龙, 吴平, 林茜, 郑佩璇, 安俊秀. (2022). Emoji 图像符号的社交功能及应用. *心理科学进展*, *30*(5), 1062–1077.]

Jones, E. E., & Pittman, T. S. (1982). Toward a general theory of strategic self-presentation. *Psychological Perspectives on the Self*, *1*(1), 231–262.

Karlan, D., & Mcconnell, M. A. (2014). Hey look at me: The effect of giving circles on giving. *Journal of Economic Behavior & Organization*, *106*, 402–412.

Kaushik, M., Singh, V., & Chakravarty, S. (2022). Experimental evidence of the effect of financial incentives and detection on dishonesty. *Scientific Reports*, *12*(1), 2680.

Kaye, L. K., Wall, H. J., & Malone, S. A. (2016). "Turn that frown upside-down": A contextual account of emoticon usage on different virtual platforms. *Computers in Human Behavior*, *60*, 463–467.

Kaye, L. K., Malone, S. A., & Wall, H. J. (2017). Emojis: Insights, affordances, and possibilities for psychological science. *Trends in Cognitive Sciences*, *21*(2), 66–68.

Kim, J., & Jeong, B. (2020). Expecting social punishment facilitates control over a decision under uncertainty by

recruiting medial prefrontal cortex. *Social Cognitive and Affective Neuroscience*, *15*(11), 1260–1270.

Kohls, G., Perino, M. T., Taylor, J. M., Madva, E. N., Cayless, S. J., Troiani, V., Price, E., Faja, S., Herrington, J. D., & Schultz, R. T. (2013). The nucleus accumbens is involved in both the pursuit of social reward and the avoidance of social punishment. *Neuropsychologia*, *51*(11), 2062–2069.

Kringelbach, M. L., & Rolls, E. T. (2003). Neural correlates of rapid reversal learning in a simple model of human social interaction. *NeuroImage*, *20*(2), 1371–1383.

Kujawa, A., Proudfit, G. H., Kessel, E. M., Dyson, M., Olino, T., & Klein, D. N. (2015). Neural reactivity to monetary rewards and losses in childhood: Longitudinal and concurrent associations with observed and self-reported positive emotionality. *Biological Psychology*, *104*, 41–47.

Leary, M. R., & Kowalski, R. M. (1990). Impression management: A literature review and two-component model. *Psychological Bulletin*, *107*(1), 34–47.

Leimar, O.,& Hammerstein, P. (2001). Evolution of cooperation through indirect reciprocity. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *268*(1468), 745–753.

Li, J., Sun, Y., Yang, Z. L., & Zhong, Y. P. (2020). Social value orientation modulates the processing of social rewards for self: Evidence from ERPs study. *Acta Psychologica Sinica*, *52*(6), 786–800.

[李珺, 孙宇, 杨子鹿, 钟毅平. (2020). 社会价值取向对自我社会奖赏加工的影响-来自 ERPs 的证据. *心理学报*, *52*(6), 786–800.]

Li, S. J., Tang, Y. Y., & Zhang, D. D. (2024). Neural mechanism of monetary and social reward processing in healthy and depressed populations. *Journal of Psychological Science*, *47*(6), 1317–1327.

[李思瑾, 汤煜尧, 张丹丹. (2024). 健康和抑郁人群金钱和社会奖赏加工的神经机制. *心理科学*, *47*(6), 1317–1327.]

Liu, C. J., & Hao, F. (2011). Social value orientation and cooperation in asymmetric social dilemmas. *Acta Psychologica Sinica*, *43*(4), 432–441.

[刘长江, 郝芳. (2011). 不对称社会困境中社会价值取向对合作的影响. *心理学报*, *43*(4), 432–441.]

Lv, F. Y., Tan, J. B., Xu, P. F., Xiong, X. L., Jin, Z. H., & Gao, D. G. (2021). The social reward and the neurocognitive mechanism of social reward processing. *Chinese Journal of Applied Psychology*, *27*(3), 189–203.

[吕房艳, 谭敬斌, 徐鹏飞, 熊希灵, 金泽慧, 高定国. (2021).社会性奖赏及其作用的神经机制. *应用心理学*, *27*(3), 189–203.]

Matyjek, M., Meliss, S., Dziobek, I., & Murayama, K. (2020). A multidimensional view on social and non-social

rewards. *Frontiers in Psychiatry*, *11*, 818.

Mazar, N., & Ariely, D. (2006). Dishonesty in everyday life and its policy implications. *Journal of Public Policy & Marketing*, *25*(1), 117–126.

Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research*, *45*(6), 633–644.

Milinski, M., Semmann, D., & Krambeck, H. J. (2002). Reputation helps solve the 'tragedy of the commons'. *Nature*, *415*(6870), 424–426.

Milinski, M., Semmann, D., Krambeck, H. J., & Marotzke, J. (2006). Stabilizing the Earth's climate is not a losing game: Supporting evidence from public goods experiments. *Proceedings of the National Academy of Sciences*, *103*(11), 3994–3998.

Mulder, L. B., Dijk, E. V., Cremer, D. D., & Wilke, H. A. M. (2006). Undermining trust and cooperation: The paradox of sanctioning systems in social dilemmas. *Journal of Experimental Social Psychology*, *42*(2), 147–162.

Murphy, R. O., Ackermann, K. A., & Handgraaf, M. (2011). Measuring social value orientation. *Judgment and Decision Making*, *6*(8), 771–781.

Nagin, D. S., & Pogarsky, G. (2003). An experimental investigation of deterrence: Cheating, self-serving bias, and impulsivity. *Criminology*, *41*(1), 167–194.

Nowak, M. A., & Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature*, *437*(7063), 1291–1298.

Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, *2*, 343.

Pfeiffer, T., Tran, L., Krumme, C., & Rand, D. G. (2012). The value of reputation. *Journal of the Royal Society Interface*, *9*(76), 2791–2797.

Ramirez-Marin, J. Y., & Shafa, S. (2018). Social rewards: The basis for collaboration in honor cultures. *Cross Cultural & Strategic Management*, *25*(1), 53–69.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Computation*, *20*(4), 873–922.

Romano, A., Balliet, D., Yamagishi, T., & Liu, J. H. (2017). Parochial trust and cooperation across 17 societies. *Proceedings of the National Academy of Sciences*, *114*(48), 12702–12707.

Rosenbaum, S. M., Billinger, S., & Stieglitz, N. (2014). Let's be honest: A review of experimental evidence of honesty and truth-telling. *Journal of Economic Psychology*, *45*, 181–196.

Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, *5*(4), 296–320.

Russell, Y. I., Call, J., & Dunbar, R. I. M. (2008). Image scoring in great apes. *Behavioural Processes*, *78*(1), 108–111.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78.

Shore, D. M., & Heerey, E. A. (2011). The value of genuine and polite smiles. *Emotion, 11*(1), 169–174.

Simpson, B., & Willer, R. (2008). Altruism and indirect reciprocity: The interaction of person and situation in prosocial behavior. *Social Psychology Quarterly*, *71*(1), 37–52.

Somerville, L. H., Jones, R. M., Ruberry, E. J., Dyke, J. P., Glover, G., & Casey, B. J. (2013). The medial prefrontal cortex and the emergence of self-conscious emotion in adolescence. *Psychological Science*, *24*(8), 1554–1562.

Sommerfeld, R. D., Krambeck, H. J., Semmann, D., & Milinski, M. (2007). Gossip as an alternative for direct observation in games of indirect reciprocity. *Proceedings of the National Academy of Sciences*, *104*(44), 17435–17440.

Spreckelmeyer, K. N., Krach, S., Kohls, G., Rademacher, L., Irmak, A., Konrad, K., ... & Gründer, G. (2009). Anticipation of monetary and social reward differently activates mesolimbic brain structures in men and women. *Social Cognitive and Affective Neuroscience*, *4*(2), 158–165.

Sperber, D., & Baumard, N. (2012). Moral reputation: An evolutionary and cognitive perspective. *Mind & Language, 27*(5), 495–518.

Steinel, W. (2015). Social value orientation and deception: Are proselfs liars? *Current Opinion in Psychology*, *6*, 211–215.

Tamir, D. I., & Mitchell, J. P. (2012). Disclosing information about the self is intrinsically rewarding. *Proceedings of the National Academy of Sciences*, *109*(21), 8038–8043.

Tamir, D. I., Zaki, J., & Mitchell, J. P. (2015). Informing others is associated with behavioral and neural signatures of value. *Journal of Experimental Psychology: General*, *144*(6), 1114–1123.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *The Quarterly Review of Biology*, *46*(1), 35–57.

Vabba, A., Porciello, G., Panasiti, M. S., & Aglioti, S. M. (2022). Interoceptive influences on the production of self-serving lies in reputation risk conditions. *International Journal of Psychophysiology*, *177*, 34–42.

Van Lange, P. A. (1999). The pursuit of joint outcomes and equality in outcomes: An integrative model of social value orientation. *Journal of Personality and Social Psychology*, *77*(2), 337–349.

Van Kleef, G. A. (2009). How emotions regulate social life: The emotions as social information (EASI) model. *Current Directions in Psychological Science*, *18*(3), 184–188.

Walther, J. B. (1992). Interpersonal effects in computer-mediated interaction: A meta-analysis of social and antisocial communication. *Human Communication Research*, *19*(4), 159–182.

Wang, D., Liu, T., & Shi, J. (2017). Development of monetary and social reward processes. *Scientific Reports*, *7*(1), 11128.

Wang, Z., Li, Q., Nie, L., & Zheng, Y. (2020). Neural dynamics of monetary and social reward processing in social anhedonia. *Social Cognitive and Affective Neuroscience*, *15*(9), 991–1003.

Wu, J., Balliet, D., & Lange, P. A. M. V. (2016). Reputation management: Why and how gossip enhances generosity. *Evolution & Human Behavior*, *37*(3), 193–201.

Yuan, B., Wang, X. P., Yin, J., & Li, W. Q. (2023). The role of cross-situational stimulus generalization in the formation of trust towards face: A perspective based on direct and observational learning. *Acta Psychologica Sinica*, *55*(7), 1099–1114.

[袁博, 王晓萍, 尹军, 李伟强. (2023). 跨情境的刺激泛化在面孔信任形成中的作用：基于直接互动与观察学习的视角. *心理学报*, *55*(7), 1099–1114.]

Zhang, Y. H., Li, H., & Wu, Y. (2020). The application of computational modelling in the studies of moral cognition. *Advances in Psychological Science*, *28*(7), 1042–1055.

[张银花, 李红, 吴寅. (2020). 计算模型在道德认知研究中的应用. *心理科学进展*, *28*(7), 1042–1055.]

Zhang, Z., Zhang, F., Yuan, S., Guo, F. B., & Wang, Y. W. (2015). Psychometric analysis of the SVO slider measure in Chinese cultural context. *Studies of Psychology and Behavio*, *13*(3), 404–409.

[张振, 张帆, 原胜, 郭丰波, 王益文. (2015). 社会价值取向滑块测验中文版的测量学分析. *心理与行为研究*, *13*(3), 404–409.]

Zhu, L., Jenkins, A. C., Set, E., Scabini, D., Knight, R. T., Chiu, P. H., King-Casas, B., & Hsu, M. (2014). Damage to dorsolateral prefrontal cortex affects tradeoffs between honesty and self-interest. *Nature Neuroscience*, *17*(10), 1319–1321.

The Influence of Social Reward and Punishment on Deception

YUAN Bo, ZHAO Jingshi, QI Dan, ZHAO Tong, HU Jiaqi

(*Department of Psychology, Ningbo University, Ningbo 315211, China*)

Abstract

Deception is a complex behavior involving the dissemination of misleading information or the intentional omission of relevant details, conveyed both verbally and nonverbally. While extensive research has explored strategies to mitigate deception, rewards and punishments have emerged as key deterrents. Prior studies have predominantly focused on tangible incentives, such as monetary or token-based rewards and punishments. However, given that deception frequently occurs in interpersonal contexts, social rewards and punishments—eliciting positive or negative emotional responses—may also influence deceptive behavior. Despite this, their precise impact remains unclear

To address this gap, the present study employs three experiments to examine the interplay between social rewards, social punishments, and deception, while also investigating the underlying mediating and moderating mechanisms. Experiment 1 engaged 30 participants in a signaling game (also known as the sender-receiver game) to assess whether social rewards and punishments influence deception in a manner similar to monetary incentives. Experiment 2, with 60 participants, extended this investigation by incorporating an adapted reputation concern scale to explore reputation concern as a potential mediator. Experiment 3 examined the moderating role of social value orientation (SVO). Participants with different SVOs, identified using the SVO slider measure, were recruited to determine whether social value orientation moderates the effects of social rewards and punishments on deception.

Findings from Experiment 1 demonstrated that social rewards and punishments, akin to monetary incentives, reduce deceptive behaviors, with social punishments being more effective than social rewards. Drift-Diffusion Modeling (DDM) analysis revealed that under conditions of social and monetary rewards and punishments, the drift rate (v) was significantly lower compared to the condition without incentives, indicating that both forms of incentives promote evidence accumulation favoring non-deceptive behavior. Experiment 2 established reputation concern as a

mediator in the relationship between social incentives and deception. Specifically, social incentives heightened reputation concerns, leading to reduced deception, whereas monetary incentives did not exert this effect. Experiment 3 identified social value orientation as a moderator: individuals with a pro-social orientation exhibited heightened reputation concerns in response to social incentives, thereby reducing deception, while this effect was absent in pro-self individuals.

Overall, this study confirms that social rewards and punishments effectively reduce deceptive behavior. By elucidating the psychological mechanisms involved and broadening the empirical understanding of social incentives, these findings offer valuable insights into mitigating deception in interpersonal interactions. Future research can further explore the moderating effects of different social contexts or individual differences to more comprehensively understand the boundary conditions of how social rewards and punishments influence deceptive behavior.

Keywords social reward and punishment, deception, reputation concern, social value orientation, Drift-Diffusion Modeling

附录 A　声誉关注量表

1、在前一部分任务的决策过程中，我会考虑对方如何评价我这个人。

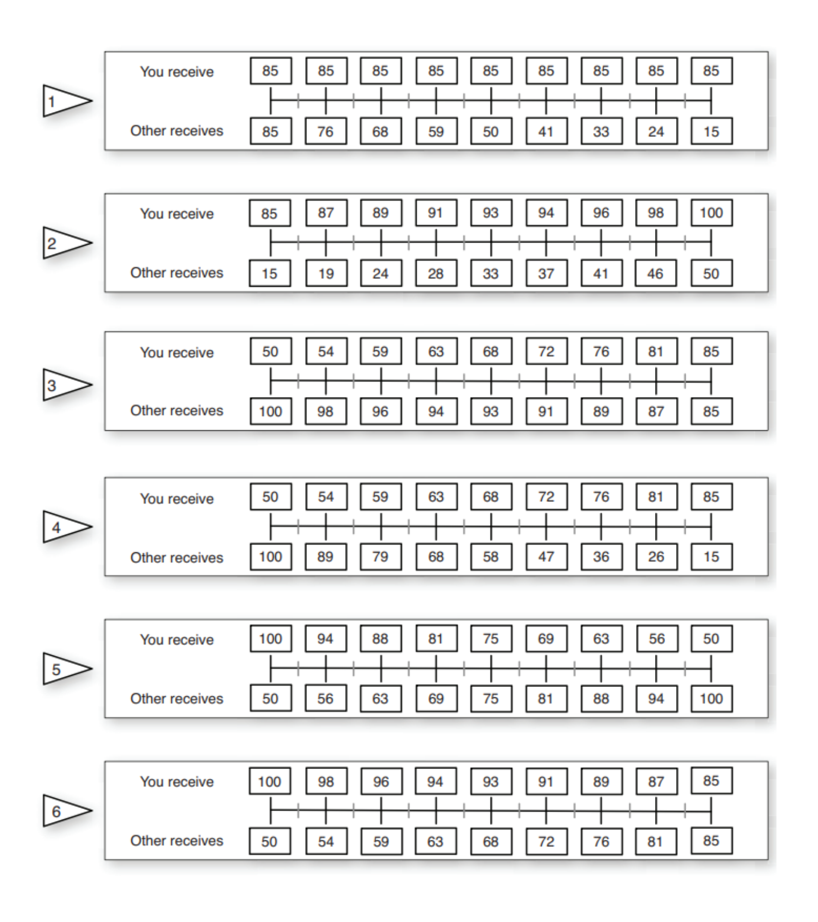
①完全不符合 ②不符合 ③比较不符合 ④有点不符合 ⑤不确定⑥ 有点符合 ⑦比较符合 ⑧符合 ⑨完全符合

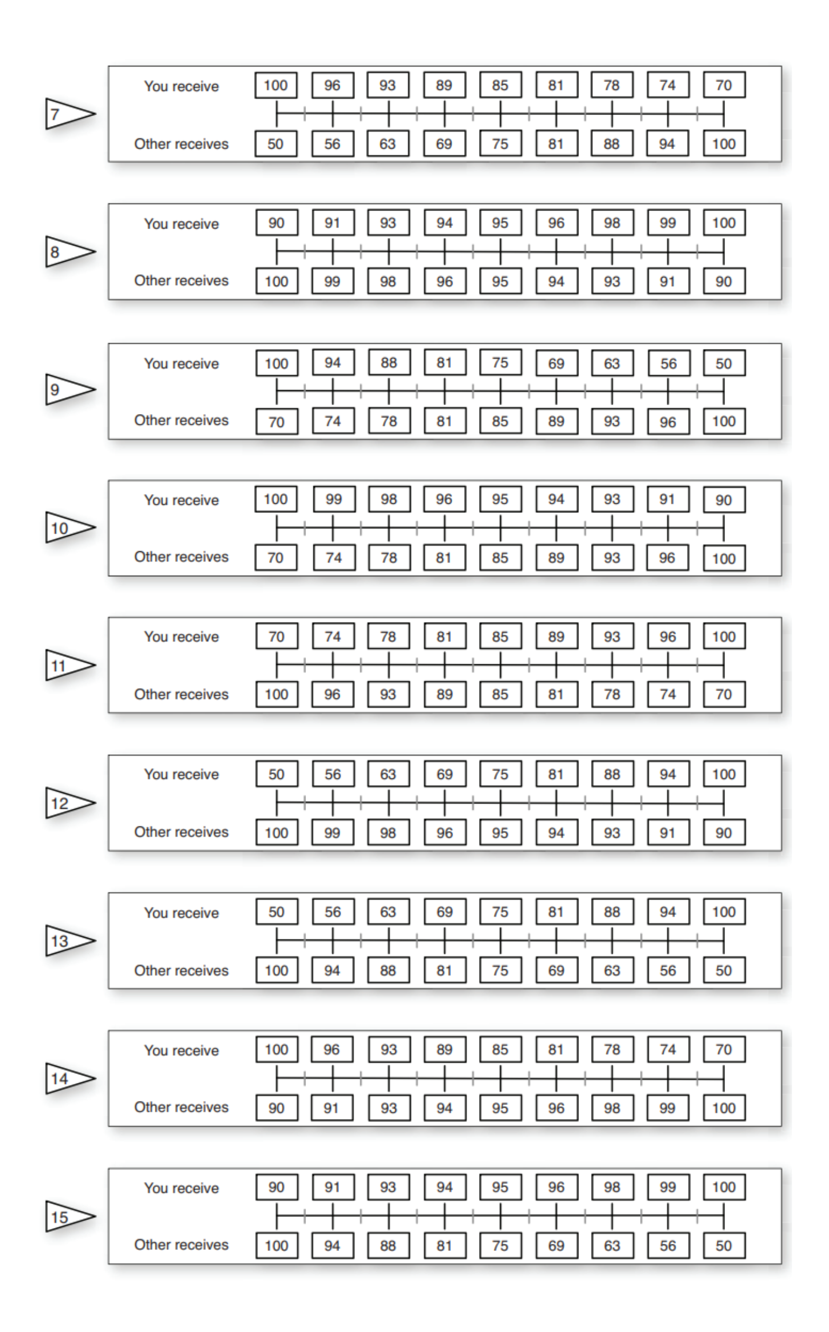
2、在前一部分任务的决策过程中，我完全不在乎对方给我的回应。

①完全不符合 ②不符合 ③比较不符合 ④有点不符合 ⑤不确定⑥ 有点符合 ⑦比较符合 ⑧符合 ⑨完全符合

附录 B　社会价值取向滑块测验

在这个任务中，请你想象自己与某个人(用 TA 表示)被随机分配在一起，你们彼此并不认识，将来也不会见面。你的所有选择将会完全保密。你将在你和 TA 之间做出一系列关于分配资源的决策。对于下面的每个问题，请选择下列收益分配图中你最偏好的结果分配。你的决策将会决定你和 TA 各自所得的金钱数。结果没有正误之分，只是反映个人偏好。




7
You receive 100 96 93 89 85 81 78 74 70
Other receives 50 56 63 69 75 81 88 94 100
8
You receive 90 91 93 94 95 96 98 99 100
Other receives 100 99 98 96 95 94 93 91 90
9
You receive 100 94 88 81 75 69 63 56 50
Other receives 70 74 78 81 85 89 93 96 100
10
You receive 100 99 98 96 95 94 93 91 90
Other receives 70 74 78 81 85 89 93 96 100
11
You receive 70 74 78 81 85 89 93 96 100
Other receives 100 96 93 89 85 81 78 74 70
12
You receive 50 56 63 69 75 81 88 94 100
Other receives 100 99 98 96 95 94 93 91 90
13
You receive 50 56 63 69 75 81 88 94 100
Other receives 100 94 88 81 75 69 63 56 50
14
You receive 100 96 93 89 85 81 78 74 70
Other receives 90 91 93 94 95 96 98 99 100
15
You receive 90 91 93 94 95 96 98 99 100
Other receives 100 94 88 81 75 69 63 56 50

附录 C　数据分析表

附表 1 实验 3 社会价值取向*奖惩类型混合设计方差分析结果

效应	*df*	*F*	*p*	η_p^2
社会价值取向	1, 191	30.99	< 0.001	0.14
奖惩类型	2, 58	40.66	< 0.001	0.183
社会价值取向 * 奖惩类型	2, 382	5.44	0.005	0.03

附表 2 实验 3 社会价值取向在奖惩类型上的描述性统计($M \pm SD$)

社会价值取向	社会奖惩	金钱奖惩	无奖惩
亲自我取向	0.60 ± 0.32	0.46 ± 0.34	0.65 ± 0.31
亲社会取向	0.32 ± 0.27	0.30 ± 0.26	0.45 ± 0.32